Optical packet switched design with relaxed maximum hardware parameters and high service-class granularity for flexible switch node dimensioning

Nord, Martin

Published in:
11th International Telecommunications Network Strategy and Planning Symposium. NETWORKS 2004

Link to article, DOI:
10.1109/NETWKS.2004.1341889

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Nord, M. (2004). Optical packet switched design with relaxed maximum hardware parameters and high service-class granularity for flexible switch node dimensioning. In *11th International Telecommunications Network Strategy and Planning Symposium. NETWORKS 2004* (pp. 453-458). IEEE. https://doi.org/10.1109/NETWKS.2004.1341889

Optical packet switch design with relaxed maximum hardware parameters and high service-class granularity for flexible switch node dimensioning

Martin Nord
Research Centre COM, Technical University of Denmark, DTU B-345V, DK-2800 Lyngby, Denmark
and with Telenor R&D, N-1331 Fornebu, Norway.

Abstract

This paper proposes a quality of service differentiation algorithm, improving the service class granularity and isolation of our recently presented waveband plane based design. The design aims at overcoming potential hardware limitations and increasing the switch node dimensioning flexibility in core networks. Exploiting the wavelength dimension for contention resolution, using partially shared wavelength converter pools, avoids optical buffers and reduces wavelength converter count. These benefits are illustrated by numerical simulations, and are highlighted in a dimensioning study with three service classes.

1 Introduction

In the context of core telecommunication networks, Optical Circuit Switching (OCS) may evolve into purely statistical multiplexed networks, such as Optical Packet- and Burst Switching (OPS/OBS) networks, in order to improve network utilisation and transfer delay. These optical statistically multiplexed networks cannot be store-and-forward networks, as opposed to electrical packet networks, since the optical equivalent of random access memory for buffering does not exist. Optical Fiber Delay Lines (FDLs) [1], or electrical buffers [2] can be used to some extent, but their number should be minimised to limit space consumption or to reduce cost associated with O/E interfaces, respectively. Similar to [3], this study aims to avoid buffers, by resolving contention purely in the wavelength dimension, thereby minimising transfer delay and avoiding packet re-ordering in the optical layer. This calls for Wavelength Converters (WCs), which transfer the data from the original WDM channel onto a new WDM channel, governed by the local laser source wavelength.

An OPS or OBS switch faces several potential physical constraints, such as an electrical control plane bottleneck, maximum switch matrix size and maximum laser tuning range. The Waveband Plane (WP) based node design was proposed in [4], and addresses three main issues in OPS design: overcoming hardware constraint, reducing cost and offering Quality of Service (QoS) differentiation. The former is achieved by relaxing maximum switch matrix size and laser tuning range through switch design parallelism.

Cost is potentially reduced by reducing the wavelength converter count, through introducing of Shared Per Waveband Plane (SPWP) TWC pools. The latter point may be realized by mapping traffic from different Classes of Service (CoS) on to different WPs, each having a Packet Loss Rate (PLR) linked to its SPWP TWC pool size. This also links the PLR of each CoS to the WP cost.

This work introduces a new WP-internal QoS differentiation, in order to increase QoS granularity and to enable low PLRs. Although the design is proposed in an OPS context, the principle is general, and may be used on OBS and OCS switch designs. Since circuit switching is preferable for many applications, hybrid networks, supporting both OCS and OBS/OPS, may offer the best overall performance/complexity ratio. Gradually introducing OBS/OPS calls for network node designs capable of supporting the migration. The WP design can achieve this, by mapping the OCS CoS onto a corresponding WP, with appropriate control plane and data-plane solution [4]. Moreover, in such a hybrid network, there may be little need for a CoS supporting very low PLRs. Client traffic with PLR requirements below e.g. 10^{-5} could instead be transported on the OCS network with zero loss.

This paper is organised as follows: Section 2 details the switch design. Section 3 details the WP-internal QoS differentiation algorithm and discusses performance aspects. Section 4 highlights how a design that combines WP-based and WP-internal QoS differentiation provides flexibility for node dimensioning, enabling reduction in important maximum hardware parameters and overall TWC count.

2 Shared Per Waveband Plane (SPWP) switch design

2.1 The Waveband Plane (WP) concept

We study a core switch node with F incoming and F outgoing fibres, each having a WDM channel count of W. There are f WPs, each denoted WP_i, which all receive packets from w_i wavelengths from each fibre, by use of a passive waveband demultiplexer, as illustrated in fig. 1 c). Consequently, WP_i switches Fw_i wavelengths. The switching in each WP is independent of the other planes, and a passive coupler realizes the recombination. Since the channels on different WPs do not overlap there is no blocking at this point. As discussed in [4], the independent waveband switching will give higher blocking than multi-stage wavelength-plane [5] and fibre-plane [1] designs with coordinated recombination. This is intuitively understood by e.g. studying the situation where w+1 packets on the same WP contend for a specific fibre, when there are free outputs to the same fibre from a parallel WP. However, compared to these designs, the SPWP design avoids a second switching stage, and decouples w from W. As shown in table 1, the WP concept thereby provides a more flexible means to overcome hardware limitations in the switch dimensioning.

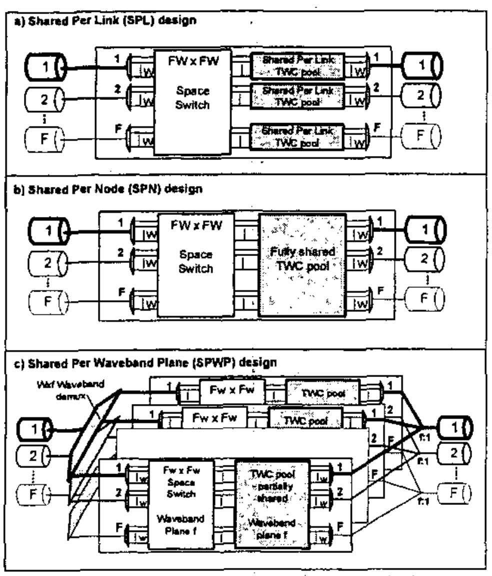

Figure 1. Generic SPL, SPN and SPWP designs.

2.2 Shared TWC pools

For a given WP without buffers and with dimension (Fw), one can resolve contention in the wavelength domain, by equipping all outputs going to the same fibre with a Fixed Output-wavelength WC (FOWC). Increasing the wavelength count dramatically lowers the PLR, as studied e.g. in [6]. This is due to economies of scale in the wavelength domain. However, since every packet in principle does not have to be wavelength converted, one may reduce WC count by only allocating WCs to contending packets. This calls for hardware modification to enable the sharing, as well as replacing WCs with Tunable WCs (TWCs). Hence, for such a solution to pay off, the WC count reduction must be significant, in order to compensate for the associated hardware complexity increase.

Fig. 1 a) and b) illustrate two main categories of TWC pool designs: Shared Per Node (SPN) [7] and Shared Per Link (SPL) [8]. The former has a single pool, shared by all packets, while the latter has a pool per output link (fibre), shared only by packets going to that link. We define the WC Ratio (WCR) as the ratio of the WC count and the number of wavelengths containing packets sharing these WCs. For a given Packet Loss Rate (PLR), economies of scale reduce the required WCR in the SPN approach. This comes at the expense of an increased TWC pool hardware complexity, compared to SPL design. The SPL pool can be implemented simply by equipping a fraction of the outputs to link with a TWC. SPN TWC pools can be implemented in a feedback design [7], by adding FW*WCR ports, each equipped with a TWC. The SPWP TWC pool can be implemented in an analogue way, with the increased freedom of varying the WCR between the WPs.

Table 1 gives important hardware parameters for SPWP designs with TWC pools in a feedback design. The port count denotes the number of outputs of the space switch, including those needed for implementing the TWC pool, and assumes that all planes have the same WCR.

Table 1. Hardware parameters.

HW parameter	Per Switch	WP size/max
Space switch port count	FW(1+WCR)	Fw*(1+WCR)
TWC count	FW*WCR	-
TWC tuning range	N/A	w
Wavelength demux	2fF	1 x w
Waveband demux	F	1 x f
Passive coupler	F	F:1

For a given F, the *overall* number of TWCs and the space switch port count is independent of the waveband size, w. On the other hand, the *maximum* TWC tuning range (expressed in number of channels), space switch port count, multiplexer size, waveband and output coupler is proportional to w. This shows that introducing a WP based concept with a waveband of dimension w (w<W), relaxes maximum switch hardware parameters.

2.3 QoS differentiation in SPWP approach

In the core node, traffic is mapped onto the WPs based on which waveband they belong to. Section 3 details how the WCR_i governs the PLR of WP_i. Hence, one can offer QoS differentiation on a network basis, by mapping traffic onto wavebands according to the required CoS, in the OPS network ingress nodes. This requires distribution of information on what PLR to be associated with each waveband in all network nodes. In this approach, the control unit of each WP does not have to inspect any CoS field in packet headers to provide the appropriate QoS. As discussed in [4], this distributed per-WP scheduling without QoS processing may alleviate a potential electronic control plane bottleneck.

To increase QoS granularity, at the same time maintaining economies of scale in the wavelength dimension, for efficient contention resolution, we introduce a WP internal QoS algorithm in the SPWP design. Although this adds to the WP internal QoS scheduling, the control plane bottleneck is still much reduced, compared to a corresponding central switch, since scheduling parallelism is maintained and since only two CoS are considered per WP.

3 Performance

3.1 WP internal QoS differentiation

As discussed e.g. in [9], several methods exist for QoS differentiation in OPS/OBS network. A main difference is that OBS enables different offset-based priority schemes, but this comes at the expense requiring reading of burst duration and tracking the start- and end time of current allocations and future reservation. In OPS, the possibility of differentiating access to WCs and wavelengths [1], as well as only WCs has been explored, e.g. in [10]. As studied below, the hereby proposed QoS differentiation algorithm is particularly suited to the WP based design, by explicitly differentiating the access both to the TWCs and to the output wavelengths.

As depicted in fig. 2, when no WP internal QoS differentiation is applied, the PLR of each WP will depend on the WCR, for a given F and w. We here consider mapping up to two different CoS on each WP, termed CoS 1 and CoS 2 in the following, by implementing the WP internal QoS differentiation algorithm. In this case, the PLR of each CoS also depends on the fraction of reserved WCs in the pool, R_{WC}, and the fraction of the reserved output wavelengths going to the same output fibre, R_{WL}.

The scheduling algorithm for the considered WP is given below. Point 1 ensures a minimum WCR usage, by preferring "direct wavelength mapping". Point 2 ensures that when no direct wavelength mapping can be made, CoS 1 packets have prioritised access to switch resources. When successful switching cannot be made, the packet in question is discarded.

1. Regardless of CoS, switch packet to fibre, if its current wavelength on that fibre is free.
2. If no direct wavelength mapping possible, resolve contention by converting the packet to a randomly chosen free output wavelength on desired fibre.
    a. If the packet belongs to CoS 1 it can use any free WC or wavelength.
    b. If the packet belongs to CoS 2 it can only be allocated if the fraction of the WP's free WCs *and* of the WP's free output wavelengths to the desired output fibre are above R_{WC} and R_{WL}, respectively.

Considering the electronic hardware complexity, the scheme requires keeping track of the WP's TWC pool usage and output wavelength usage. First, a table look-up is made to see if its own wavelength is free. If this is not the case, the scheduler checks whether the packet still can be allocated depending to the packet's CoS, by checking the status of a wavelength and a WC counter. If both resources are available, both the WC and output wavelength are allocated by first-fit incremental searches among the potential wavelengths and wavelength converters. The complexity is hence moderate, compared to offset-based or advanced FDL buffer algorithms, which calls taking timing information into account in the allocation.

We denote the PLR of the higher and lower priority CoS on each WP, PLR_{CoS1} and PLR_{CoS2}, respectively. Their isolation is denoted I_{PLR}, here defined as $I_{PLR} = (PLR_{CoS2}/PLR_{CoS1})$. The scheduling algorithm has the particularity that CoS 2 packets are accepted, when the corresponding wavelength at the output fibre is free, even when either the fraction of free WCs or output wavelengths are below R_{WC} and R_{WL}, respectively. Whilst this makes sense from a pure WC utilisation point of view, since the packet does not consume any WCs anyway, the effect is more complex with respect to wavelength utilisation. This direct wavelength mapping preference decreases PLR_{CoS2}, and the increased utilisation may cause rejection of some CoS 1 packets, increasing PLR_{CoS1}. Hence I_{PLR} decreases, but so does also $PLR_{OVERALL}$, particularly when WCs are scarce. An interesting aspect is that positive discrimination of direct mappings increases the chance that the input wavelength of a later arriving packet from the same input fibre is among the free wavelengths on the same output fibre, effectively increasing the chance of another direct mapping. This effect increases with decreasing w and F, and will be subject to a more thorough investigation, out of the scope of this study.

3.3 Numerical simulations

3.3.1 Input parameters

Event-driven simulations are used to evaluate the switch performance. We consider symmetric input traffic, i.e. that the desired output of each packet is uniformly distributed among the output fibres. We focus on a node adjacency, F, of 4, representative of a mesh network. Each input wavelength has an independent generator, which generates packets with exponential distributed duration, according to a Poisson arrival process. These packets are clocked out on each wavelength on a FIFO basis. The switch performance is measured in terms of PLR. The 95 % confidence intervals, obtained by 10 successive simulations with different random parameters are indicated. We study the PLR range between 10^{-5} and 10^{-1}.

3.3.2 No WP internal QoS differentiation

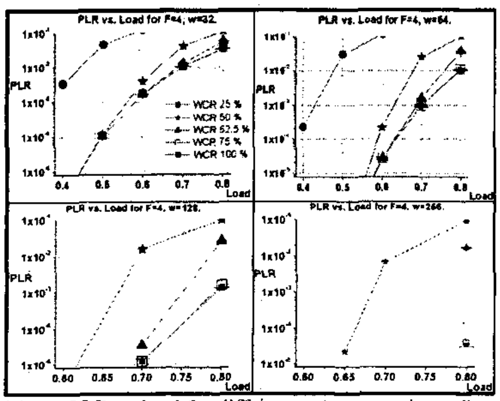

Fig. 2. PLR vs load for different w (common legend).

Fig. 2 shows how the PLR depends on load for different w ranging from 32 to 256, and for WCRs in the range of of 0.25 to 1. Our simulation results show how decreasing the WCR from 1 induces a PLR penalty, when the load is sufficiently high. To enable practically feasible node dimensioning, a fixed channel load is beneficial. Since high utilisation is a main motivation for introducing statistical multiplexing, we aim at a reasonably high channel load of 0.7, maintained throughout this study.

Fig. 3 shows PLR vs. WCR, for different waveband sizes. It indicates that this channel load enables PLRs from 1 % and to below 10^{-5}, for waveband sizes of 32 to 128, for a sufficiently high WCR. These cases have a WCR threshold at a value slightly below the channel load, above which PLR does not decrease further. Reducing WCRs from 1 to 0.6 reduces the PLR by factors between 1.5 and 15, for w between 32 and 128.

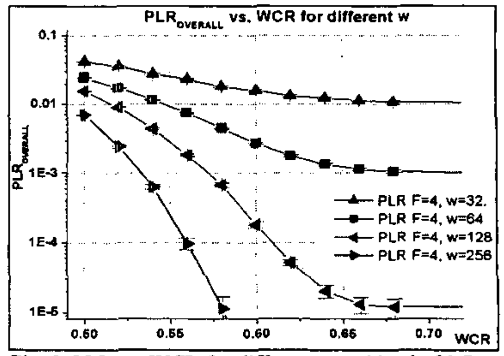

Fig. 3. PLR vs. WCR for different w and load of 0.7

3.3.3 WP internal QoS differentiation

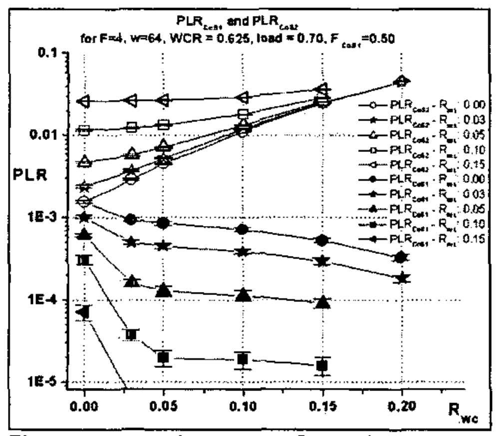

Fig. 4. PLR_{CoS1} and PLR_{CoS1} vs. R_{WC} and R_{WL}

When employing WP internal QoS differentiation, we aim at offering a very low PLR_{CoS1}, on the expense of an increased PLR_{CoS2}. We denote the fraction of CoS 1 traffic, with respect to the sum of both CoS, F_{CoS1}, and set it to 0.5 in this study. Fig. 4 shows how the PLRs evolve with R_{WC} and R_{WL}, for w=64, F_{CoS1}=0.5 and WCR=62.5 %. Let us first consider the two axis of isolation independently, by only reserving either wavelengths or WCs. For R_{WL}=0, increasing R_{WC} from 0 to 0.2 increases PLR_{CoS2} to 4.6 %, whilst PLR_{CoS1} decreases to around 3.2×10^{-4}. Conversely, for R_{WC}=0, increasing R_{WL} from 0 to 0.15 increases PLR_{CoS2} to 2.6 %, whilst PLR_{CoS1} decreases to around 7.2×10^{-5}.

Note that the algorithm's explicit reservation of a fraction of both wavelengths and wavelength converters increases its efficiency: for a combination of R_{WC} and R_{WL} that has a lower PLR_{CoS1} than the two above cases, PLR_{CoS2} is also decreased. Put another way: $PLR_{OVERALL}$ is reduced for the same I_{PLR}. For the dimensioning of a switch with a given w, one must adjust WCR, R_{WC} and R_{WL} in order to obtain the required PLR_{CoS1} and PLR_{CoS2}. Similar simulations have been carried out for different parameters, and are used for the dimensioning study in the next section.

4 Switch Dimensioning

In 4.1.1 and 4.1.2 we detail the traffic matrix and the strategies used for node dimensioning. A fully optimised dimensioning study is out of the scope of this paper. It is e.g. possible to study different traffic matrices, increase the number of WP internal QoS levels, choose another network global load or study variations on the CoS mapping strategies studied here. Nevertheless, this study illustrates the main properties of this design, as shown by the results, given in 4.1.3, and further discussed in 4.1.4.

4.1.1 Traffic Matrix

Table 2 shows the assumed traffic matrix, defining the PLR requirements and relative traffic load for three CoS, termed "GOLD", "SILVER" and "BRONZE". We consider two overall WDM channel counts, W=128 and W=256, compatible with demonstrated transmission capacities using 10 Gbit/s channels [11]. For F=4 and load=0.7, this corresponds to offered traffic of 3.6 and 7.2 Tbit/s, respectively.

Table 2. Traffic Matrix for the dimensioning study

CoS	Ratio	CoS PLR threshold
GOLD	F_{GOLD} =25 %	$PLR_{MAX,GOLD}=2x10^{-5}$
SILVER	F_{SILVER} =50 %	$PLR_{MAX,SILVER}=2x10^{-3}$
BRONZE	F_{BRONZE} =25 %	$PLR_{MAX,BRONZE}=2x10^{-2}$

4.1.2 Dimensioning strategies

The dimensioning strategies investigated are listed below. Note that the "SPN" strategy is equivalent to a SPWP with a single plane and thus w=W. The "SPWP" corresponds to the pure SPWP design presented in [4], and that the "SPWP+" uses the WP internal QoS differentiation algorithm discussed above.

- **SPN** : Map all traffic onto WP_1. Find a WCR that matches the most demanding PLR threshold.
- **SPWP** : Map all CoS onto three corresponding WPs, termed WP_1-WP_3. Find a WCR for each WP that respect the corresponding PLR threshold.
- **SPWP+**: Map the GOLD and BRONZE CoS on to the same WP, termed WP_1. Choose a WCR_1, R_{WC} and R_{WL} that respect both the PLR thresholds. Map the SILVER CoS onto WP_2, with a suitable WCR_2.

4.1.3 Results

The WP number, waveband size and WCR resulting from each strategy are given in each row of table 3. The corresponding overall WCR, WCR_{TOTAL}, is given in the rightmost column. When the PLR threshold of a CoS cannot be reached, this is marked with "–" in the corresponding column and the WCR_{TOTAL} is omitted.

Table 3. Results of the dimensioning study.

Strategy \ CoS	GOLD (WP_i, w_i, WCR_i)	SILVER (WP_i, w_i, WCR_i)	BRONZE (WP_i, w_i, WCR_i)	WCR_{TOTAL}
Switch dimensioning for W=128				
SPN	WP_1, 128, 64%	WP_1, 128, 64%	WP_1, 128, 64%	64 %
SPWP	-	WP_2, 64, 62%	WP_3, 32, 58%	-
SPWP+	WP_1, 64, 60%	WP_2, 64, 62%	WP_1, 64, 60%	61 %
Switch dimensioning for W=256				
SPN	WP_1, 256, 56%	WP_1, 256, 56%	WP_1, 256, 56%	56 %
SPWP	-	WP_2, 128, 56%	WP_3, 64, 52%	-
SPWP+	WP_1, 128, 55%	WP_2, 128, 56%	WP_1, 128, 55%	55.5 %

For W=128, the SPN strategy enables reaching PLR_{GOLD} for a WCR of 64%. With the pure SPWP Strategy, PLR_{GOLD} cannot be respected. The SPWP+ strategy enables all PLR thresholds, by choosing R_{WC}= 0.06 and R_{WL} = 0.12, and it has the lowest overall TWC count.

For W=256, the overall WCR decreased, but the results are quite similar to above. The SPWP+ strategy enables all PLR thresholds, by choosing R_{WC}= 0.03 and R_{WL} = 0.05.

4.1.4 Discussion

The dimensioning study shows that the pure SPWP strategy is not compatible with simultaneously having high CoS granularity and CoS with very low PLR thresholds. For the studied traffic matrix, the SPN and SPWP+ strategies enables total WC savings from around 45 % to 36 %, compared to using a dedicated FOWC per wavelength.

Due to increased economy of scale, all strategies perform better when W increases. For a given W and fixed PLR thresholds, the SPN strategy is insensitive to the fraction of traffic from each CoS. The pure SPWP strategy benefits from an increased fraction of the most demanding CoS, since this increases the wavelength count of the WP containing the most critical CoS. The SPWP+ strategy, on the other hand, benefits from a decrease in the fraction of CoS_{GOLD} relative to CoS_{BRONZE}.

4.2 Hardware maximum parameters

Table 4. Switch hardware parameters.

Strategy \ W	WC count 128	WC count 256	Max Tuning range 128	Max Tuning range 256	Max port count 128	Max port count 256
FOWC	512	1024	0	0	512	1024
SPN	328	573	128	256	840	1597
SPWP+	313	568	64	128	415	799

Table 4 sums up hardware maximum parameters of each of the designs resulting from the different strategies, based on the formulas in 2.2. We omit the SPWP strategy, since it did not reach the most demanding PLR threshold for the studied traffic matrix. The FOWC design corresponds to a single switch, with a FOWC at each switch matrix port output.

For W=256, the SPN strategy gives around 64 % to 100 % increase in maximum switching size, compared to FOWC and the SPWP+ strategy, respectively. Compared to the latter, the SPN strategy requires a doubling of the laser tuning range, measured in number of achievable channels.

The SPWP+ strategy enables a 22 % reduction in maximum switch matrix size and a 45 % reduction in WC count, compared to the single plane FOWC design. This comes at the expense of an increased processing complexity in the WP employing the QoS differentiation scheme, the use of tuneable lasers and an overall increase in the switch matrix port count, corresponding to the $WCR_{OVERALL}$.

5 Conclusion

The recently proposed OPS/OBS design, employing TWC pools that are Shared Per Waveband Plane, helps overcoming hardware constraints. Moreover, the design may significantly reduce the overall number of wavelength converters, compared to designs with a dedicated fixed output wavelength converter. This comes at the expense of replacing the WCs by TWCs, and by increasing overall switch matrix port counts correspondingly.

The waveband plane internal QoS differentiation algorithm proposed in this study, decreases overall loss for the same CoS isolation, compared to QoS algorithms that only reserves wavelengths or wavelength converters. Employing this algorithm in the Shared Per Waveband Plane design improves dimensioning flexibility. We highlighted how this can be exploited to increase CoS granularity and to offer low packet loss rates, without resorting to buffers, at reasonably high wavelength loads.

These features, together with its compatibility for a hybrid circuit- and packet switched network, make the design a promising optical element candidate, when facing hardware scalability constraints in a multiple service class optical network environment.

Acknowledgements

Thanks to V. B. Iversen at the Research Center COM, for fruitful discussions. The author acknowledges the Research Council of Norway for Ph.D. funding.

References

[1] F. Callegati et al., "Exploitation of DWDM for Optical Packet Switching With Quality of Service Guarantees", IEEE JSAC 20 (1), pp. 190-201, Jan. 2002.

[2] S. Bjornstad et al., "A scalable optical packet switch for variable length packets employing shared electronic buffering", Proc. 28th European Conference on Optical Communication (ECOC), P4.07, Copenhagen, 2002.

[3] S. L. Danielsen et al., "WDM packet switch architectures and analysis of the influence of tuneable wavelength converters on the performance," IEEE JLT, vol. 15, pp. 219–227, Feb. 1997.

[4] M. Nord, "Waveband based multi-plane optical packet switch with partially shared wavelength converters", Proc. 8th Optical Networking Design and Modelling (ONDM) Conference, Ghent, 2004.

[5] M. J. O'Mahoney et al., "An optical packet switched network (WASPNET)– concept and realization", in Optical Networks Magazine, pp. 46-53, November/December 2001.

[6] J. S. Turner, "WDM Burst Switching for Petabit Data Networks", Proc. 25th Optical Fiber Communication (OFC) conference, vol. 2, pp. 47-49, 2000.

[7] V. Eramo, M. Listanti, "Packet Loss in a Bufferless Optical WDM Switch Employing Shared Tunable Wavelength Converters", in IEEE JLT 18 (12), pp. 1818- 1833, Dec. 2000.

[8] D. Li, X. Jia, "Placement of Wavelength Converters in Shared-per-Link Structure in WDM Networks", Proc. International Conference on Parallel Processing Workshops (ICCP Workshops), pp. 277-282, 2001.

[9] M. Nord et al., "OPS or OBS in the core network?" Proc. 7th Optical Networking Design and Modelling (ONDM) Conference, pp. 179-198, Budapest, 2003.

[10] H. Overby, N. Stol, "A Teletraffic Model for Service Differentiation in OPS networks" Proc. 8th Optoelectronic and Communications Conference (OECC), vol. 2, pp.677-678, 2003.

[10] C. Hullin et al. "Ultra long-haul 2500 km terrestrial transmission of 320 channels at 10 Gbit/s over C+L bands with 25 GHz wavelength spacing", Proc. 28th European Conference on Optical Communication (ECOC), 1.1.3, Copenhagen, 2002.